

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2002, 00:09:10 ; Search time 8498.8 seconds
(without alignments)
30.345 Million cell updates/sec

Title: US-09-851-670-3

Perfect score: 24

Sequence: 1 tggctgctgctgctgctgctgcaag 24

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 111874

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_estlin:*
3: em_estlin:*
4: em_estlin:*
5: em_estlin:*
6: em_estlin:*
7: em_estlin:*
8: em_estlin:*
9: em_estlin:*
10: em_estlin:*
11: em_estlin:*
12: em_estlin:*
13: em_estlin:*
14: em_estlin:*
15: em_estlin:*
16: em_estlin:*
17: em_estlin:*
18: em_estlin:*
19: em_estlin:*
20: em_estlin:*
21: em_estlin:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	15	62.5	59	11	BG237930
2	14	58.3	49	11	BG237930
3	14	58.3	49	11	BG237930
4	13.6	56.7	45	13	AZ314853
5	13.6	56.7	45	13	AZ314853
6	13.4	55.8	43	10	AA482116
7	13.4	55.8	46	13	AZ448066
8	13.4	55.8	50	10	AU102540
9	13.4	55.8	50	10	AU104358
10	13.4	55.8	53	10	AA413642
11	13.4	55.8	56	13	AZ445458
12	13.2	55.0	33	13	AZ341288

C 13	13.2	55.0	34	13	CNS000EXA	AL069767 Drosophila
C 14	13	54.2	33	13	AZ350538	AZ350538 1M088B08
C 15	13	54.2	42	13	AZ349062	AZ349062 1M088B08
C 16	13	54.2	52	10	AA209172	AA209172 z64a03.r
C 17	12.8	53.3	46	11	A1790825	A1790825 UK28d11.Y
C 18	12.8	53.3	49	11	H84362	H84362 YV85C09.r1
C 19	12.8	53.3	50	10	AU102537	AU102537 AU102537
C 20	12.8	53.3	50	10	AU102538	AU102538 AU102538
C 21	12.8	53.3	50	10	AU102539	AU102539 AU102539
C 22	12.8	53.3	50	10	AU102541	AU102541 AU102541
C 23	12.8	53.3	50	10	AU102543	AU102543 AU102543
C 24	12.8	53.3	50	10	AU102544	AU102544 AU102544
C 25	12.8	53.3	50	10	AU102546	AU102546 AU102546
C 26	12.8	53.3	50	10	AU102547	AU102547 AU102547
C 27	12.8	53.3	50	10	AU103070	AU103070 AU103070
C 28	12.8	53.3	50	10	AU104596	AU104596 AU104596
C 29	12.8	53.3	50	10	AU104603	AU104603 AU104603
C 30	12.8	53.3	50	10	AU104604	AU104604 AU104604
C 31	12.8	53.3	50	10	AU104611	AU104611 AU104611
C 32	12.8	53.3	50	10	AU104612	AU104612 AU104612
C 33	12.8	53.3	50	10	AU104613	AU104613 AU104613
C 34	12.8	53.3	50	10	AU104616	AU104616 AU104616
C 35	12.8	53.3	50	10	AU104617	AU104617 AU104617
C 36	12.8	53.3	50	10	AU104618	AU104618 AU104618
C 37	12.8	53.3	50	10	AU104620	AU104620 AU104620
C 38	12.8	53.3	50	10	AU104626	AU104626 AU104626
C 39	12.8	53.3	50	10	AU104628	AU104628 AU104628
C 40	12.8	53.3	50	10	AU104632	AU104632 AU104632
C 41	12.8	53.3	50	10	AU104633	AU104633 AU104633
C 42	12.8	53.3	50	10	AU104635	AU104635 AU104635
C 43	12.8	53.3	50	10	AU104636	AU104636 AU104636
C 44	12.8	53.3	50	10	AU104642	AU104642 AU104642
C 45	12.8	53.3	50	10	AU104643	AU104643 AU104643

ALIGNMENTS

RESULT 1
BG237930 59 bp mRNA EST 13-FEB-2001
LOCUS sab09d06.y1 6m-cl071 glycine max cdna clone GENOME SYSTEMS CLONE
DEFINITION ID: Gm-cl071-948 5', mRNA sequence.
ACCESSION BG237930.1 GI:12773003
VERSION
KEYWORDS
SOURCE
ORGANISM soybean.
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE
AUTHORS Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Holla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, N., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.

TITLE
JOURNAL
COMMENT
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R./Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estevats@wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com

```

Insert Length: 1349      Std Error: 0.000
Seq primer: Promega -21m13
High quality sequence stop: 1.
      Location/Qualifiers
          1. 49

```

```

/organism="Homo sapiens"
/db_xref="GDB:3782543"
/db_xref="taxon:9606"
/clone="IMAGE:209702"
/clone_lib="Soares fetal liver spleen 1nFLS"

```

```

/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACGTGACAGATTAAATTAAGACTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fátima Bonaldo."

```

	10 a	16 c	17 g	3 f
Parity	58.3%	Score 14;	DB	
73.9% Pred	No	5.5a		

Query Match	58.3%	Score 14	DB 11	Length 49
Best Local Similarity	73.9%	Pred. No. 5	5e+04	
Matches 17	Conservative 0	Mismatches 6	Indels 0	Gaps 0
OY	2	gagctgctctgagatctcggaag	24	
db	44	GGGTCGCTCTGGACGCGTCFNAAG	22	

QY	2	gactggtcttgatgtcgaag	24
Db	44	GGCTGCTCTGGACGCTNAAG	22

RESULT 3
EE371806

LOCUS	BE371806	59 bp	mRNA	EST	21-JUL-2000
DEFINITION	601217793.F1	NCI_CGAP_Lu29	Mus musculus	CDNA	IMAGE:3586678 5',
ACCESSION	BE371806				
VERSION	BE371806.1	GI:9312723			
KEYWORDS	EST.				
ORGANISM	house mouse.				
SOURCE	Mus musculus				

ACKNOWLEDGMENTS
We thank the following for their helpful comments on this manuscript: Eukariyota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus 1 (bases 1 to 59)
REFERENCE
NIH-MGC <http://mgc.ncl.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9agab3-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLB)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLB at: <http://image.jnl.gov>

```

FEATURES
source
1. 59
/location/Qualifiers
/organism="Mus musculus"
/strain="CZECH II (fetal)"
/db_xref="taxon:10090"
/clone="IMAGE:3586678"
/clone_lib="NCI CGAP_Lu29"
/tissue_type="Spontaneous tumor, metastatic to mammary
stem cell origin."
/lab_host="DHI08"
/note="Organ: lung; Vector: pCMV-SPORT6; Site: 1; Salt:

```


ORGANISM Homo sapiens
COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 43)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.
TITLE Washu-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. 43
/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone="IMAGE:756406"
/clone.lib="Soares ovary tumor NBHOT"
/sex="Female"
/tissue="ovarian tumor"
/lab.host="DH10B (ampicillin resistant)"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a modified polylinker; Site: 1: Not I; Site: 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACATCTGAGAGGAGGAGGCGGGGTTTTTTTTTTTTTTT 3'] double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 9 a 13 c 11 g 10 t
ORIGIN

Query Match 55.8%; Score 13.4; DB 10; Length 43;
Best Local Similarity 73.9%; Pred. No. 9.4e+04;
Matches 17; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ggcgtgctggagatcggaag 24
||||| ||||| | |||||
Db 38 GCGTCTCTGGAACCTCTCAAG 16

RESULT 7
LOCUS AZ448066 46 bp DNA GSS 04-OCT-2000
DEFINITION IM0245106r Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0245106 R, DNA sequence.
ACCESSION AZ448066
VERSION AZ448066.1 GI:10600489
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 46)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A., and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 1000 Std Error: 0.00
Plate: 0245 row: 1 column: 06
Seq primer: CACACAGCAACAGCTATGACC
Class: Plasmid ends
High quality sequence stop: 46.
Location/Qualifiers
1. 46
/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UUGC1M0245106"
/clone.lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab.host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (9114732114[9b]AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 7 a 4 c 23 g 12 t
ORIGIN

Query Match 55.8%; Score 13.4; DB 13; Length 46;
Best Local Similarity 73.9%; Pred. No. 9.5e+04;
Matches 17; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ggcgtgctggagatcggaag 24
||||| ||||| | |||||
Db 14 GCGTGGGTGGGTGGGTGGGAGC 36

RESULT 8
LOCUS AU102540 50 bp mRNA EST 05-APR-2001
DEFINITION AU102540 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HRC05567, mRNA sequence.
ACCESSION AU102540
VERSION AU102540.1 GI:13552061
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki, Y., Tanoda, T., Taira, H., Mizushima-Sugano, J., Sese, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Morishita, S., Okubo, K., Suyama, A., and Sugano, S.
TITLE Fine structural analysis of transcription start sites of human mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL Unpublished (2001)

COMMENT

Contact: Yutaka Suzuki
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp
 Suzuki.Y., Yoshitomo-Nakagawa.K., Maruyama.K., Suyama.A. and Sugano
 S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

Location/Qualifiers
 1. 50

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HRC05567"

BASE COUNT 7 a 11 c 25 g 7 t

ORIGIN

Query Match 55.8%; Score 13.4; DB 10; Length 50;
 Best Local Similarity 73.9%; Pred. No. 9.5e+04;
 Matches 17; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 2 tggctgctggagatgctggaag 24
 ||| || ||||| ||| ||
 Db 19 GGCAGGTTGGCATGCGCGCTGG 41

RESULT 9

AU104358 50 bp mRNA EST 05-APR-2001
 LOCUS AU104358 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 DEFINITION HBMA260082, mRNA sequence.

ACCESSION AU104358
 VERSION AU104358.1 GI:13553879
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 50)
 Suzuki.Y., Tsunoda.T., Taira.H., Mizushima-Sugano.J., Seese.J., Hata
 H., Ota.T., Isogai.T., Tanaka.T., Nakamura.Y., Morishita.S., Okubo
 K., Suyama.A. and Sugano.S.
 Fine Structural analysis of transcription start sites of human

TITLE

MRNAs using full-length enriched and 5'-end enriched cDNA libraries
 unpublished (2001)

JOURNAL

Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp

Suzuki.Y., Yoshitomo-Nakagawa.K., Maruyama.K., Suyama.A. and Sugano
 S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source

1. 50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HBMA260082"
 /clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT

8 a 8 c 26 g 6 t 2 others

ORIGIN

Query Match 55.8%; Score 13.4; DB 10; Length 50;
 Best Local Similarity 70.8%; Pred. No. 9.5e+04;
 Matches 17; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 tggctgctggagatgctggaag 24
 ||| ||||| ||| ||
 Db 26 TGGCNGTGGGCGTGTGCTGGG 49

RESULT 10

AA413642

LOCUS AA413642 53 bp mRNA EST 04-AUG-1997
 DEFINITION vc57b03.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
 IMAGE:778637 5' similar to gb:U16580 Mouse mRNA for
 peptidylarginine deaminase, complete cds (MOUSE);, mRNA sequence.

ACCESSION

AA413642.1 GI:2073779

VERSION

AA413642.1

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murine; Mus.
 1 (bases 1 to 53)
 Marra.M., Hillier.L., Allen.M., Bowles.M., Dietrich.N., Dubuque.T.,
 Geisel.S., Kucaba.T., Lacy.M., Le.M., Martin.J., Morris.M.,
 Schellenberg.K., Steptoe.M., Tan.F., Underwood.K., Moore.B.,
 Theising.B., Wylie.T., Lennon.G., Soares.B., Wilson.R. and
 Waterston.R.

The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

COMMENT

Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:471493
 Seq primer: -40ml3 fwd. EF from Amersham.

FEATURES

source

1. 53
 /organism="Mus musculus"
 /strain="B6D2 F1/J"
 /db_xref="taxon:10090"
 /clone="IMAGE:778637"
 /clone_lib="Knowles Solter mouse 2 cell"
 /tissue_type="embryo"
 /dev_stage="2-cell"
 /lab_host="DH10B"
 /note="Organ: embryo; Vector: Bluescribe (modified);
 Site:1: Multi; Site:2: SalI; Cloned unidirectionally from
 mRNA prepared from 13,500 2-cell stage embryos. Primer:
 SalI(dT): 5'-CGGTCGACCGTCGACCGTCTTTTCTTTT-3'.
 were cloned into the MluI/SalI sites of a modified
 Bluescribe vector using commercial linkers (NEB).
 Average insert size: 1.2 kb."

BASE COUNT 10 a 12 c 15 g 16 t

ORIGIN

Query Match 55.8%; Score 13.4; DB 10; Length 53;
 Best Local Similarity 73.9%; Pred. No. 9.5e+04;
 Matches 17; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 tggctgctggagatgctggaag 23
 ||| ||||| |||| ||||

Db 29 TGCCTCTCTCTGTGTCTGANG 51

RESULT 11

AZ445458

LOCUS AZ445458 56 bp DNA GSS 04-OCT-2000
 DEFINITION 1M0241J14F mouse 10kb plasmid U06C1M library Mus musculus genomic
 clone U06C1M0241J14 F, DNA sequence.

ACCESSION

AZ445458.1 GI:10595300

VERSION

AZ445458.1

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
AUTHORS	1 (bases 1 to 56) Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamli,C., Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly, M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL	Unpublished (2000)
COMMENT	Contact: Robert B. Weiss University of Utah Genome Center Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert length: 10000 Std Error: 0.00 Plate: 0241 row: 3 column: 14 Seq primer: CGTTGTAAACGACGCCACGT Class: plasmid High quality sequence stop: 56.
FEATURES	location/Qualifiers 1..56 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UUCG1M0241J14" /clone_lib="Mouse 10kb plasmid UUCG1M library" /sex="Male" /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-" /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1147321149b1ar129072.1)' a copy-number inducible derivative of plasmid RL. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT	16 a 23 c 2 g 15 t
ORIGIN	
Query Match	55.8%; Score 13.4; DB 13; Length 56;
Best Local Similarity	73.9%; Pred. No. 9.5e+04;
Matches 17; Conservative	0; Mismatches 6; Indels 0; Gaps 0;
Oy	2 ggcctgctggatctcggaag 24
Db	37 gGTTGTAATGTGATGAGAGAG 15
RESULT 12	
AZ341288	33 bp DNA GSS 29-SEP-2000
LOCUS	1M0073P05R Mouse 10kb plasmid UUCG1M library Mus musculus genomic
DEFINITION	AZ341288
ACCESSION	AZ341288.1 GI:10417390
VERSION	GSS.
KEYWORDS	house mouse.
SOURCE	Mus musculus
ORGANISM	

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	I (bases 1 to 33)			
TITLE	Dunn,D., Loyacque,A., Barber,M., Baecorn,T., Duval,B., Hamll,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.			
JOURNAL	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts			
COMMENT	Unpublished (2000) Contact: Robert B. Weiss University of Utah Genome Center Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0073 row: P column: 05 Seq primer: CACACAGGAACACCTATGACC Class: plasmid ends High quality sequence stop: 33.			
FEATURES	Location/Qualifiers 1..33 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UUCGM0073P05" /clone_lib="Mouse 10kb plasmid UUCGM library" /sex="Male" /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-" /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g114732114[gb AF129072.1]), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."			
BASE COUNT	7 a 1 c 14 g 11 t			
ORIGIN				
Query Match	55.0% Score 13.2; DB 13; Length 33;			
Best Local Similarity	83.3% Pred. No. 1.1e+05;			
Matches 15; Conservative	0; Mismatches 3; Indels 0; Gaps 0;			
QY	2 ggcctgctcggagatcg 19 			
Db	7 GGCTGCTGTAGAGATGTTG 24			
RESULT 13				
CNS00EXA/c	CNS00EXA 34 bp DNA GSS 04-JUN-1999			
LOCUS	Drosophila melanogaster genome survey sequence TET3 end of BAC:			
DEFINITION	BACR29D20 of RPCI-98 library from Drosophila melanogaster (fruit fly) genomic survey sequence.			
ACCESSION	AL069767			
VERSION	AL069767.1 GI:4949926			
KEYWORDS	GSS.			

SOURCE
ORGANISM fruit fly.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS 1 (bases 1 to 34)
TITLE Genoscope.
JOURNAL Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammaster in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..34
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR29D20"
/note="end : TET3"
Location/Qualifiers

BASE COUNT
ORIGIN 6 a 18 c 4 g 0 t 6 others

Query Match 55.0% Score 13.2; DB 13; Length 34;
Best Local Similarity 78.9% Pred. No. 1.le+05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 tgctggtctggtgcatctg 19
||| ||| ||| ||| ||| |||
Db 21 TGGGGGCTCGCGGNTGTCG 3

RESULT 14
AZ350538/c 33 bp DNA GSS 29-SEP-2000
LOCUS
DEFINITION IM0088E08F Mouse 10kb plasmid UNGC1M library Mus musculus genomic
clone UNGC1M008E08 F, DNA sequence.
ACCESSION
AZ350538
VERSION
AZ350538.1 GI:10429775
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 33)
REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0088 row: E column: 08
Seq primer: CGTTGTAAACGACGGCCACT
Class: plasmid ends
High quality sequence stop: 33.
Location/Qualifiers

FEATURES
source
1..33
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UNG1M008E08"
/clone_lib="Mouse 10kb plasmid UNGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g1147321149b1af29072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN 5 a 14 c 1 g 13 t

Query Match 54.2% Score 13; DB 13; Length 33;
Best Local Similarity 76.2% Pred. No. 1.3e+05;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ggcgtgctgggcatctgga 22
||| ||| ||| ||| ||| |||
Db 23 GGATGCTCGGAGGGGAAA 3

RESULT 15
AZ349062/c 42 bp DNA GSS 29-SEP-2000
LOCUS
DEFINITION IM0085H22R Mouse 10kb plasmid UNGC1M library Mus musculus genomic
clone UNGC1M0085H22 R, DNA sequence.
ACCESSION
AZ349062
VERSION
AZ349062.1 GI:10428299
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 42)
REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606

Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 1000 Std Error: 0.00
Plate: 0085 row: H column: 22
Seq primer: CACACAGGAAACGCTATGACC
Class: plasmid ends
High quality sequence stop: 42.

FEATURES
source

1. 42
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC1M0085H22"
/clone_lib="Mouse 10kb plasmid UUC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

9 a 16 c 4 g 13 t

Query Match 54.2%; Score 13; DB 13; Length 42;
Best Local Similarity 76.2%; Pred. 1.4e+05;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 tggctggtctggagatcgga 21
||| |||| ||| |||| ||
Db 28 TGGTTGATGCAATCTGAGA 8

Search completed: March 9, 2002, 00:09:11
Job time: 11027 sec